Journal of Medical Virology 59:52–59 (1999)

Molecular Characterization of the 5′ Non-Coding Region of South African GBV-C/HGV Isolates: Major Deletion and Evidence for a Fourth Genotype

Timothy J. Tucker,[1,2]* Heidi Smuts,[2] Peter Eickhaus,[1] Simon C. Robson,[1] and Ralph E. Kirsch[1]

[1]MRC/UCT Liver Research Center, Department of Medicine, University of Cape Town, Cape Town, South Africa
[2]Department of Medical Microbiology/SAIMR, University of Cape Town, Cape Town, South Africa

GB virus C/hepatitis G virus (GBV-C/HGV) has been characterised as a novel flavivirus, and to date three known genotypes have been cloned. Greater genetic variation of GBV-C/HGV has been demonstrated in West African isolates, but no major deletions have been shown in the 5′ non-coding region (NCR). The 5′NCR regulates protein translation via an internal ribosomal entry site (IRES). We cloned, sequenced, and analysed a 344-bp polymerase chain reaction (PCR) product, representing >60% of the 5′NCR, from 32 GBV-C/HGV PCR-positive volunteers. Wild-type virus amplicons were detected in all samples. However, 5/32 (15.6%) also amplified another fragment of between 205 and 231 bp. Sequence analysis showed all cloned PCR fragments to be GBV-C/HGV-specific. A typical deletion of 113–131 bp with minor variation was detected in isolates generating the smaller bands. RNA secondary structure analysis showed the deletions to be over domains II and III. This finding suggests that nucleotides 303–444 may be non-essential for 5′NCR functioning. Phylogenetic analysis demonstrated a novel fourth South African genotype, distinct from genotypes 1–3 with DNA distances of >0.1000. Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) values for the wild-type and mutant samples were normal. This study documents the first major deletion in the 5′NCR of GBV-C/HGV, and suggests that bases 303–444 may not be essential for viral replication and ribosomal entry. A fourth GBV-C/HGV genotype appears to predominate in South Africa. ***J. Med. Virol. 59:52–59, 1999.***

KEY WORDS: GBV-C/HGV; 5′NCR; genotype; mutation; South Africa

INTRODUCTION

The GB virus-C and hepatitis G virus (both hereinafter GBV-C/HGV) have been isolated recently, and characterised as novel variants of the same transfusion-transmissible, positive strand, RNA flavivirus, initially reported to be associated with hepatitis [Simons et al., 1995; Yoshiba et al., 1995; Heringlake et al., 1996; Linnen et al., 1996]. Although the precise link between GBV-C/HGV infection and disease is uncertain [Alter, 1996], current evidence does not support its association with liver disease [Alter H, et al., 1997; Alter M, et al., 1997], and no other disease entity has been consistently associated with infection with any variant of GBV-C/HGV. Data from Africa and South America show a greater GBV-C/HGV prevalence than that seen in developed countries [Dawson et al., 1996; Tucker et al., 1997; Bassit et al., 1998], and recent analysis has suggested that GBV-C/HGV may be of African origin [Tanaka et al., 1998]. Phylogenetic analyses of both coding and non-coding regions demonstrate the presence of three genotypes, which show consistent geographical clustering with West Africa, USA/Europe, and Asia [Muerhoff et al., 1996b; Muerhoff et al., 1997; Mukaide et al., 1997; Okamoto et al., 1997; Katayama et al., 1998].

Much of the current understanding of GBV-C/HGV biology is based on prior knowledge of the molecular features of the hepatitis C virus (HCV) and other flaviviruses, due to their similar genomic organization [Leary et al., 1996]. The most conserved regions of both GBV-C/HGV and HCV are their 5′ non-coding regions (5′NCR), making these the most sensitive targets for viral detection by polymerase chain reaction (PCR) [Muerhoff et al., 1996a; Mukaide et al., 1997]. The extreme 5′ portion of the 5′NCR of GBV-C/HGV has been described [Hsieh et al., 1997]. Although this group proposed a 465-nucleotide non-coding region, subsequent evidence suggests that the GBV-C/HGV start codon may be 163 nucleotides further downstream of this area [Pickering et al., 1997; Katayama et al., 1998].

*Correspondence to: Dr. Timothy J. Tucker, Dept. of Medical Microbiology, University of Cape Town Medical School, Observatory 7925, Cape Town, South Africa. E-mail: tim@medmicro.uct.ac.za

Accepted 11 January 1999

Although a 36-nucleotide insertion has recently been demonstrated in the NS5a region of Central African GBV-C/HGV variants [Tanaka et al., 1998], no major deletions have been shown to date in the 5′NCR of any GBV-C/HGV, hepatitis C virus (HCV), GBV-B, or GBV-A isolate.

The 5′NCR of flaviviruses are thought to regulate translation of a single downstream open reading frame that codes for both functional and structural proteins [Brown et al., 1992; Tsukiyama-Kohara et al., 1992; Muerhoff et al., 1995; Linnen et al., 1996; Simons et al., 1996]. Experimental evidence suggests that there is no GBV-C/HGV sequence-specific ribosome scanning mechanism, but rather an internal ribosome entry site (IRES) that controls ribosome docking and the initiation of translation [Simons et al., 1996]. The secondary (and tertiary) structure of the 3′ terminal portion of the 5′NCR RNA, as well as the initial portion of the coding region, appears important for ribosome binding in flaviviruses [Simons et al., 1996]. Fukushi et al. [1994] have documented that a complete 5′NCR of HCV is required as an internal initiation site for ribosomal entry. The highly conserved nature and similarity of the GBV-C/HGV and HCV 5′NCR suggests the same may hold for GBV-C/HGV protein translation.

No isolates from Southern Africa have been included to date in the studies characterising the 5′NCR region of GBV-C/HGV. Thus, this study aimed to elucidate the nature of the 5′NCR of South African variants of GBV-C/HGV, and documents the first major deletions in this important regulatory region. In addition, evidence supporting the presence of a fourth GBV-C/HGV genotype is presented.

MATERIALS AND METHODS

Specimens

Thirty two GBV-C/HGV RNA positive plasma specimens were obtained from volunteers recruited in the Western and Eastern Cape Provinces of South Africa, as described previously [Tucker et al., 1997]. The blood samples were collected from individuals residing in both rural villages and a metropolitan urban area. Positive and negative control sera were supplied by Boehringer Mannheim (Mannheim, Germany).

RNA Extraction and cDNA Synthesis

RNA was extracted from 200 μl serum using Total RNA Isolation Reagent (Advanced Biotechnologies, London, UK). The RNA was precipitated in isopropanol overnight at –20°C, pelleted, washed twice with 75% ethanol, and resuspended in 20 μl diethyl pyrocarbonate treated ultra pure water. A cDNA synthesis step was performed in a 20-μl solution containing 10 μl of the extracted RNA with final concentrations of 50 nM random hexamers, 200 μM PCR nucleotide mix, 1 U RNase inhibitor (all Boehringer Mannheim, GmbH Biochemica, Mannheim, Germany), 1× reverse transcription (RT) buffer [50 mM Tris-HCl (pH 8.3), 3 mM $MgCl_2$, 75 mM KCl] 10 mM dithiothreitol and 200 U Maloney murine leukaemia virus reverse transcriptase (Gibco BRL, Life Technologies, Scotland, UK). This reaction was incubated for 1 hr at 37°C.

GBV-C/HGV PCR and Detection

A nested PCR was performed on all samples using the following primers: sense AGGTGGTGGATGGGTGAT; antisense TGCCACCCGCCCTCACCCGAA; sense nested TGGTAGGTCGTAAATCCCGGT; and antisense nested GGAGCTGGGTGGCCCCATGCAT [Jarvis et al., 1996]. The final PCR product spanned a 344-base pair (bp) region (inclusive of nested primers) from base 139 to 483 of the complete 5′NCR (accession number U76893) [Hsieh et al., 1997]. Samples were subjected to 35 amplification cycles of 94°C, 50°C, and 72°C, each for 30 sec using a 9600 Perkin Elmer thermal cycler and thin-walled 250 μl PCR tubes (both Perkin-Elmer, Norwalk, CT). Five microlitres of the cDNA solution (or outer product) were added to a 45-μl master mix containing (final concentrations) 200 μM dATP, dCTP, dGTP, and dTTP each, 2.6 U Expand High Fidelity polymerase, and 1× PCR buffer [10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$ (pH 8.3)] (all Boehringer Mannheim) and the relevant forward and reverse primers. The PCR products were separated by electrophoresis in 2% agarose gel and visualised under ultraviolet light using ethidium bromide staining.

Southern Blot, Cloning, and Sequencing

The nested PCR fragments were confirmed as being GBV-C/HGV cDNA by Southern blot hybridization. A 186-bp, digoxigenin-labelled cDNA probe of the 5′NCR was used in the detection system described previously [Schlueter et al., 1996]. PCR products were cloned using the pCR-Script Amp SK^+ cloning kit (Stratagene, La Jolla, CA). Transformed *Escherichia coli* cells were grown overnight on agarose 2× yeast-tryptone broth containing 100 μM ampicillin, 100 μM X-gal, and 0.5 μM isopropyl-β-D-thiogalactoside (IPTG). Bacterial colonies were selected according to blue/white differentiation and grown for 16 hr in 100 ml 2× yeast-tryptone broth containing 100 μM ampicillin. Plasmid DNA was extracted (Nucleobond, Macherey-Nagel GmbH, Düren, Switzerland) and restricted with the enzyme Pvu II (Boehringer Mannheim) to confirm the inclusion of the desired cDNA fragment. All clones shown to have a smaller inserted fragment, and a selection of those with the predicted size fragment, were sequenced. Sequencing was carried out in both the forward and reverse directions using the Pharmacia ALFexpress automated sequencer (AM version 3.01, Pharmacia Biotech). Any sequence discrepancies were resolved by repeat sequencing.

For serum samples found to contain a virus with a deletion (see results below), the complete reverse transcription (RT)-PCR process was repeated for confirmation of the amplicon size. In further reactions, the RNA suspensions were heat denatured at 95°C for 5 min followed by immediate cooling on ice before RT, and PCR denaturing and annealing temperatures were varied. In addition, a separate supplementary nested

PCR was performed on samples ZASq13, ZAN11, and ZAF87 using a primer within the deleted region (sequence 5′-AGAGAGACATTGAAGGGCGACG -3′). This nested PCR was carried out for final confirmation of the presence (or absence) of the wild-type virus in these three samples, by ensuring that the presence of the wild type was not masked by the mutant outcompeting the larger fragment in the PCR process.

Sequence Analysis

The South African 5′NCR sequences were compared phylogenetically with others representing the three GBV-C/HGV genotypes described previously. The sequences used for comparison were (accession numbers): U59518-21, U59529-33, U59543-53, U76893, D87251, D87708-14, and D90601. The original HGV isolate, PNF2161 (U44402) and the GBV-C isolate (U36330) were also included in the analysis.

The nucleotide numbering used below was derived from either: (a) the position within the PCR product or (b) according to the numbering system of accession number U76893 [Hsieh et al., 1997]. Alignment of sequences was conducted using ClustalW version 1.6 (European Molecular Biology Laboratory, Heidelberg, Germany) with hand alignment by eye. Phylogenetic analysis was performed using ClustalW and Treecon for Windows version 1.1 (Antwerpen, Belgium), using the neighbour joining method and Kimura algorithm, with gaps included. Phylogenetic trees were visualised using the Treeview computer program version 1.2 (IBLS, University of Glasgow). ClustalW was used to assess the statistical reliability of the tree using the Felsenstein bootstrap method based on 100 repetitions. Intra- and inter-genotype DNA distance values were generated with ClustalW. The tree was initially generated comparing the South African wild-type variants only with the other published sequences. Deletion isolates were then included individually (using the same methodology) with the deletion gap considered to be one mutational event. The RNA secondary structure was predicted using the M-Fold computer program [Zuker et al., 1991].

Biochemistry

Serum ALT and AST were measured using conventional methods.

RESULTS

Twenty seven of the 32 samples (84.3%) amplified the predicted 344-bp PCR fragment (Fig. 1). Of the remaining five samples, 3/32 (9.4%) generated smaller PCR fragments of 214, 215, and 231 bp, respectively (ZAF87, ZAN11, and ZASq13), and 2/32 (6.3%) (ZAF3 and ZAF115) each generated two PCR bands: one of the predicted fragment size of 344 bp and another of 205 and 211 bp, respectively (Fig. 1). Southern blot analysis (after nested PCR) showed that both the predicted 344-bp PCR fragments as well as the smaller fragments were GBV-C/HGV specific. However, the presence of the wild-type cDNA fragments was not demonstrated by Southern blot in samples ZAF87, ZAN11, nor ZASq13. Restriction fragment analysis of clones confirmed the inserted PCR fragments' sizes. Many clones from each sample were analysed, and all clones from the samples with a single PCR band contained only one cDNA insert size (whether the predicted size or smaller). However, different clones from samples ZAF3 and ZAF115 contained either the 344-bp fragment or the smaller fragments. The samples consistently amplified the same amplicons, and manipulating the RT and PCR conditions did not alter this outcome. However, wild-type virus was detectable only in ZAF87, ZAN11, and ZASq13 in the supplementary nested PCR reaction (using the primer within the deleted region).

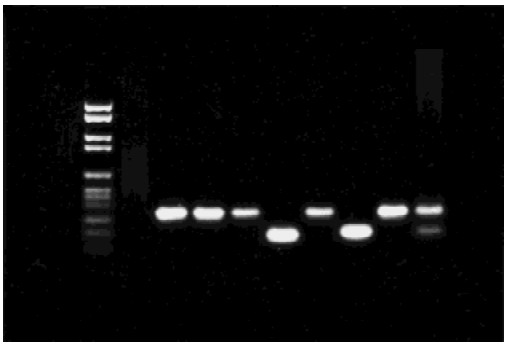
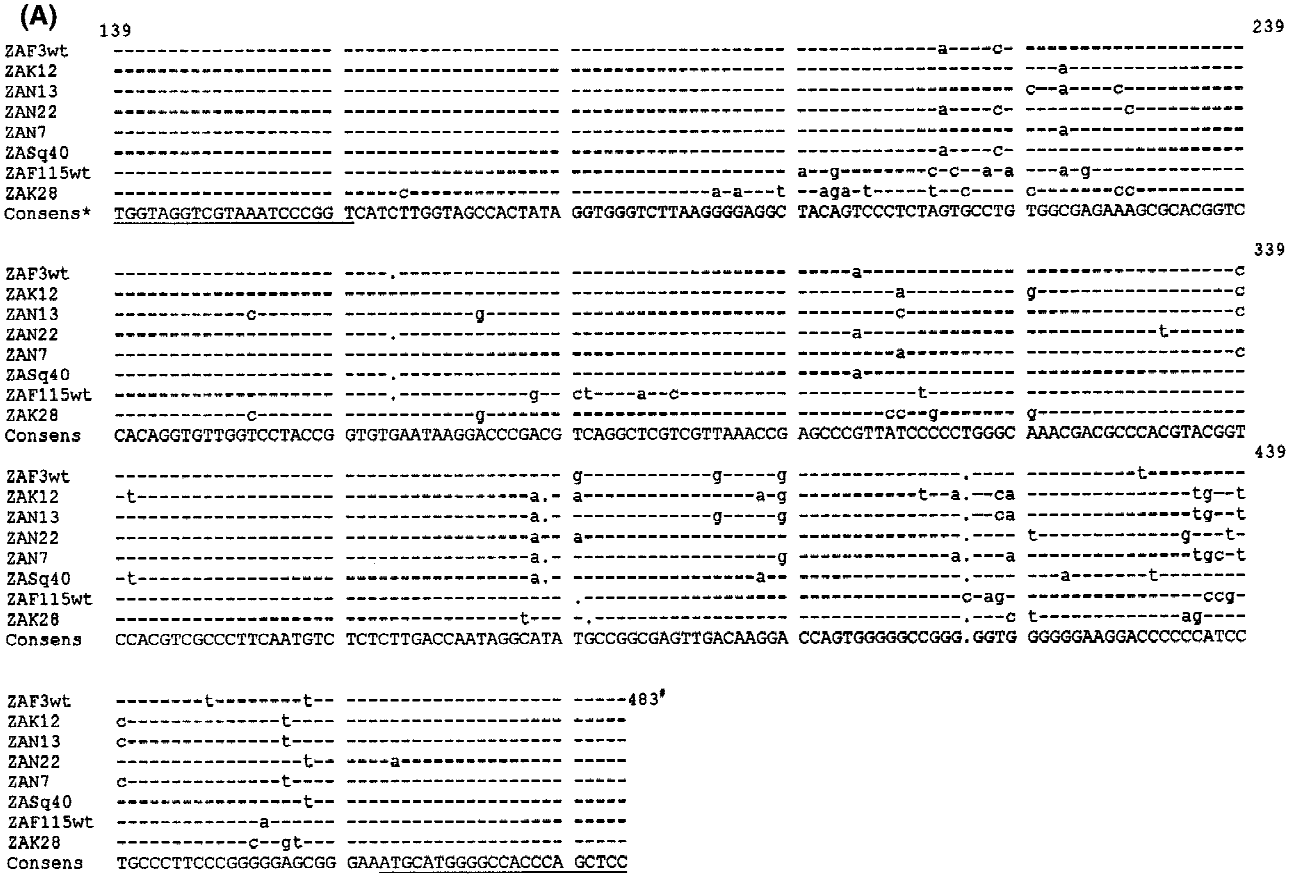


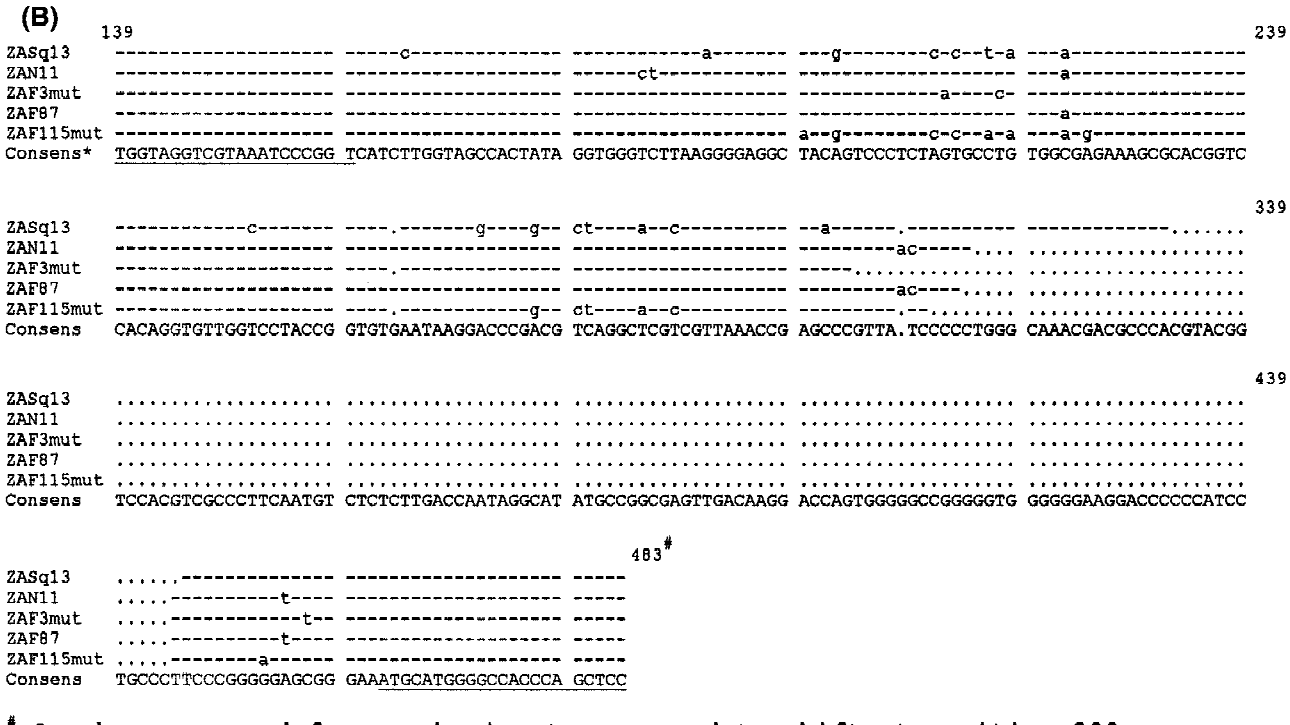
Fig. 1. Analysis of nested GB virus C/hepatitis G virus (GBV-C/HGV) polymerase chain reaction products showed either the predicted 344-bp fragment or the smaller fragment representing the deletion isolates that run below that of the positive control. Lanes 1–10 contain the following samples (in order): molecular weight marker VI (Boehringer Mannheim), negative control, positive control, ZAN22, ZAN7, ZAF87, ZAK28, ZASq13, ZASq40, and ZAF115.

Sequence analysis of the cDNA clones showed all to be GBV-C/HGV specific. The sequence alignment of eight wild-type GBV-C/HGV isolates, including the wild-type (wt) sequence of ZAF3 and ZAF115, is shown in Figure 2A. Deletant sequences from clones ZAN11, ZASq13, and ZASq87, as well as those clones containing the smaller fragment of ZAF3 and ZAF115 (designated ZAF3mut and ZAF115mut) are shown in Figure 2B. All deletions were over the same nucleotide region, with minor variation. Although the deletions ended almost uniformly at base 305 or 306 (of the predicted PCR product), each started at a different position over a span of 28 bases. Thus, using the numbering system of accession number U76893 [Hsieh et al., 1997], the deletion in clone ZASq13 was from base pair position 333 to 443. The deletions in clone ZAN11, ZAF87, ZAF3, and ZAF115 all terminated at base position 444, but started at nucleotide 314, 313, 303, and 310, respectively (Fig. 2B). The deleted sequence was over a region of the 5′NCR that was well conserved initially and then terminated after variable region III [Hsieh et al., 1997].

Phylogenetic analysis, as shown by the unrooted tree (Fig. 3), showed 9/13 (69%) of the South African isolates to cluster into a separate, distinct group when compared with isolates from West Africa, Asia, and Europe/USA. The South African cluster was not an artifact induced by the deletants, as the same phylogenetic



One base removed from numbering to accommodate shift at position 414.



One base removed from numbering to accommodate shift at position 308.

Fig. 2. The 5′ non-coding region (5′NCR) sequence alignment for South African GB virus C/hepatitis G virus (GBV-C/HGV) isolates demonstrating **(A)** the wild-type isolates amplifying the 344 bp, and **(B)** the deletion isolates. *Consens, the consensus sequence indicates the base that occurs most frequently at that position. The consensus sequence in (B) is the same as in (A). The first nucleotide of the polymerase chain reaction products corresponds to position 139 of the complete 5′NCR (accession number U76893). Dashes represent homology with the consensus sequence, and deletions are represented by periods. Forward and reverse nested primer sequences are underlined.

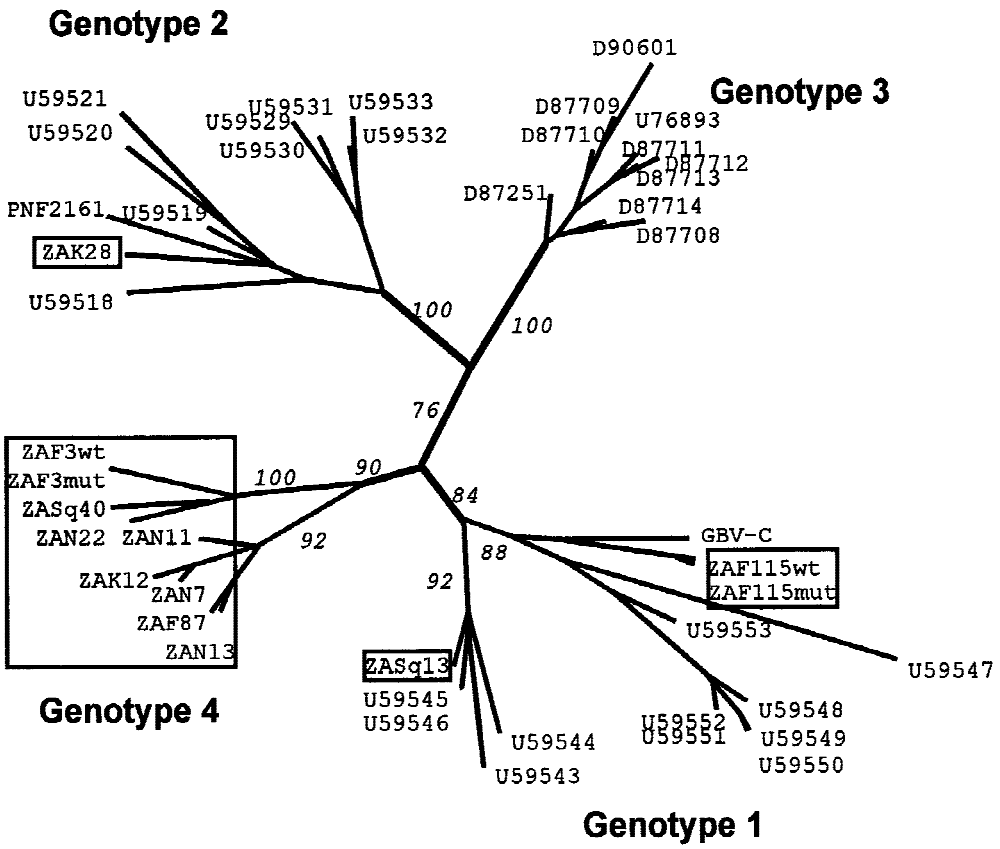


Fig. 3. Phylogenetic tree showing the DNA distance relationships between the known GB virus C/hepatitis G virus (GBV-C/HGV) genotypes 1–3 as well as the additional clustering of South African isolates into a novel fourth genotype. The branch lengths are a measure of the evolutionary distances. Bootstrap analysis values for the major branches are shown in italics.

relationships were present when the analysis was performed without the inclusion of the deletant isolates. The tree is consistent with previous phylogenetic analyses demonstrating geographical clustering of genotypes 1–3 with West Africa, Europe/USA, and Asia, respectively [Muerhoff et al., 1996b; Mukaide et al., 1997]. The statistical reliability of the tree was confirmed by bootstrap analysis (Fig. 3). The mean DNA distance values within and between the three known genotypes and the South African cluster are shown in Table I. The DNA distances between the South African group and the other genotypes were all >0.1000, that is, as large as that described previously between genotypes 1–3 [Muerhoff et al., 1996b], thus demonstrating the presence of a novel fourth genotype in South Africa.

Four South African isolates did not fall within the new South African clustering (Fig. 3). Sample ZAK28 was closely associated with the HGV-like isolates (genotype 2), whereas the samples ZASq13, ZAF115wt, and ZAF115mut were related to the GBV-C-like variants, described previously in West Africa (genotype 1). Three of the deletion isolates (ZAN11, ZAF3, and ZAF87), clustered with the new South African group, whereas ZASq13 and ZAF115mut fell into genotype 1. The different viruses (wild-type and deletion) contained in the samples ZAF3 and ZAF115 were shown to be very closely associated, suggesting that the deletion isolates may have been formed within the host.

RNA secondary structure analysis (Fig. 4) showed the predicted 5′NCR region of the wild type isolates to have a structure comparable to that described previously [Pickering et al., 1997]. The deletion involved two major stem loop structures (domains II and III) and an additional smaller loop structure. The start positions of the deletions are marked S1–S5, and the termination positions are marked T1 and T2 in Figure 4. Comparative analysis of the nucleotide sequence in Figure 2A and the predicted RNA structure in Figure 4 showed 85% of the nucleotide changes in the variable regions (II and III) to be covariant substitutions, thus allowing for the maintenance of the stem loop structure.

ALT and AST values for all 32 wild-type and mutant GBV-C/HGV samples were within the normal limits and showed no differences by statistical analysis.

DISCUSSION

The 5′NCR of flaviviruses have been extensively investigated due to their conserved nature and the putative role in regulation of translation. Although minor variation in the 5′NCRs is well established [Muerhoff et al., 1996a; Hsieh et al., 1997; Mukaide et al., 1997], this study documents the first major deletion found in the 5′NCR of certain GBV-C/HGV isolates. Our data suggest that the bases between 303 and 444 may be

TABLE I. Mean (SD) DNA Distances Between and Within GBV-C/HGV Subtypes, Derived Using ClustalW Program

	n	1	2	3	ZA
1	12	0.0677 (0.0349)	0.1200 (0.0101)	0.1109 (0.0181)	0.1126 (0.0217)
2	10		0.05139 (0.0137)	0.0941 (0.0122)	0.1055 (0.0113)
3	10			0.0274 (0.0088)	0.1113 (0.0089)
ZA	13				0.0543 (0.0210)

GBV-C, GB virus C; HGV, hepatitis G virus; n = the number of sequences analysed in each group.
The differences demonstrated the presence of a novel fourth genotype in South Africa that differs from each of genotypes 1, 2, and 3 by >0.1000.

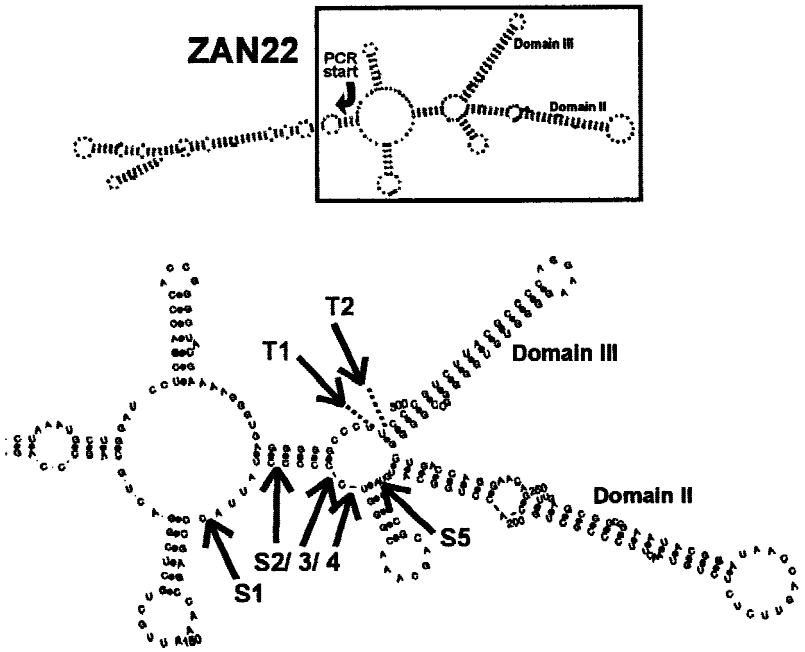


Fig. 4. Predicted RNA secondary structure of ZAN22 (inset), as generated by the Mfold program. Below is a magnification of the domains II and III, which are deleted in the five deletion isolates. All deletions terminate at base 443 or 444 (T1 and T2), and each of the deletion start positions vary, as shown by S1–S5.

non-essential for viral regulation and IRES preservation. In addition, data are presented demonstrating a fourth genotype of GBV-C/HGV that appears to predominate in South Africa.

The ability to accommodate such a large 5′NCR deletion differentiates between GBV-C/HGV and HCV, in that HCV has been shown to require the 5′NCR to be complete for replication efficiency [Fukushi et al., 1994]. It has been postulated that there is no sequence-specific motif for ribosome scanning, but rather that there is a RNA structure that acts as an IRES within the 5′NCR (and possibly the initial portion of the coding region) that facilitates ribosomal binding and subsequent translation [Simons et al., 1996]. IRES regions are found within a short distance of the ATG start codon [Brown et al., 1992]. Our data suggest that the IRES is well downstream of GBV-C/HGV base 444, so as to preserve IRES structure.

The presence of these deletions over domains II and III supports the proposal that the 5′NCR is 556 bases in length [Pickering et al., 1997; Katayama et al., 1998] and not 465 bases [Hsieh et al., 1997]. The South African deletions remove domains II and III and end only 25 bases upstream of the proposed ATG codon at base 465, thereby almost certainly altering the RNA secondary structure of the region. However, if the true start position is 93 bases further [Pickering et al., 1997; Katayama et al., 1998], the preserved region downstream of the deletion would be extended to 118 bases and thereby possibly keep domains IV and V sufficiently unaltered to preserve IRES function.

We have attempted to exclude the possibility that these deletions are due to RT or PCR errors. All serum samples showed consistent results with multiple repeat testing, and heat denaturation of the RNA immediately prior to RT did not alter this outcome. In addition, changing the PCR denaturation and annealing conditions did not alter the findings. Had the deletion findings been due to an RT artifact, more than 15% of the samples would have shown these deletions. The minor variation seen in the deletion length also suggests a more dynamic biological role involved that requires further investigation.

We were not able to detect the mutant virus in the absence of the wild type. We attempted to detect the presence of small quantities of the wild-type virus in

ZAF87, ZAN11, and ZASq13 by three methods, of which only one was successful. In the first instance, a nested PCR assay followed by Southern blot hybridization was unable to detect wild-type RNA transcripts. Secondly, purifying many clones from each of these samples and performing restriction fragment length analysis revealed none of the predicted length cDNA fragments. The wild-type transcripts were only detected in samples ZAF87, ZAN11, and ZASq13 when using a primer internal to the deletion. This finding suggests that the mutant cDNA out-competed the wild-type cDNA during the PCR process, possibly due to a vast difference in RNA titre and/or a more efficient amplification due to size difference.

The sequence data show regions of conservation within the 5′NCR as well as others with greater variability. Although the sequence data do not include the first of the GBV-C/HGV variable regions [Hsieh et al., 1997], variable regions 2 and 3 are present. Although the variability is significant in these two areas, the predicted secondary structure of the RNA shows approximately 85% of the variation to comprise covariant substitutions, thus facilitating preservation of the RNA secondary structure. Of note, all five of the deletions end just downstream of the region of greatest variability in our sequence. It is unclear whether this variability is of biological significance in the development of these deletions.

Multiple phylogenetic analyses have confirmed the presence of three GBV-C/HGV genotypes [Muerhoff et al., 1996b; Mukaide et al., 1997; Okamoto et al., 1997; Katayama et al., 1998]. However, the only isolates from Africa included in these studies were of West African origin. Our data are strongly supportive of the phylogenetic data to date, but also add a fourth genotype from Southern Africa. The DNA distance between the South African cluster and the other genotypes is as great (or greater) than the distances demonstrated between genotypes 1–3, that is, >0.1000 (Table I). The South African genotype is a less heterogenous group than genotype 1. However, this description of additional heterogeneity in GBV-C/HGV African isolates supports the hypothesis that GBV-C/HGV may originate in Africa [Tanaka et al., 1998]. Southern Africa has a predominant HCV genotype that is not common elsewhere [Smuts and Kannemeyer, 1995]. These data provide the first evidence to suggest the same for GBV-C/HGV. As South Africa is a heterogenous society with strong trade and cultural roots in both Africa and Europe/USA, it is not surprising that a small percentage of the isolates are related to genotypes that predominate outside of the region. The finding that the deletion isolates do not form a separate cluster is noteworthy and suggestive of the capacity of multiple genotypes to form this mutation. The selective advantage (if any) of this mutant genetic form has yet to be elucidated. Individuals containing both the wild type and mutants were not limited to genotype 4 (Fig. 3). For ZAF3 and ZAF115, the closeness of the two isolates in each of these individuals is suggestive of the deletion being generated after infection with a single strain, and not being due to separate infections. It remains unclear whether the mutant has the capacity to produce infectious particles, and thus whether individuals can be infected with the deletion isolate alone. Further studies will investigate this possibility.

The deletion occurs over an area proposed initially as being suitable for primer design [Muerhoff et al., 1996a]. Thus, it is possible that groups using primers over this region have missed deletion isolates. Most evidence to date suggests that GBV-C/HGV neither replicates in the liver nor causes hepatitis [Alter H, et al., 1997; Alter M, et al., 1997; Laskus et al., 1997a, 1997b]. All serum samples investigated in this study had ALT and AST levels within the normal ranges.

Thus, this study documents the molecular features of the 5′NCR of South African isolates of GBV-C/HGV and demonstrates a novel fourth genotype that appears to predominate. In addition, we demonstrate the presence of a major deletion in this important regulatory region of GBV-C/HGV, suggesting that this portion of the 5′NCR may not be essential for regulation of translation.

ACKNOWLEDGMENTS

We thank Drs. A. and C. Williamson of the University of Cape Town Department of Medical Microbiology for assistance with the preparation of this manuscript, and Ms. D. James for her technical assistance with the generation of sequence data.

REFERENCES

Alter HJ. 1996. The cloning and clinical implications of HGV and HGBV-C. N Engl J Med 334:1536–1537.

Alter HJ, Nakatsuji Y, Melpolder J, Wages J, Wesley R, Wai-Kuo Shih J, Kim JP. 1997. The incidence of transfusion-associated hepatitis G virus infection and its relation to liver disease. N Engl J Med 336:747–755.

Alter MJ, Gallagher M, Morris TT, Moyer LA, Meeks EL, Krawczynski K, Kim JP, Margolis HS. 1997. Acute non A-E hepatitis in the United States and the role of hepatitis G virus infection. N Engl J Med 336:741–746.

Bassit L, Kleter B, Ribeiro-dos-Santos G, Maertens G, Sabino E, Chamone D, Quint W, Saez-Alquezar A. 1998. Hepatitis G virus: prevalence and sequence analysis in blood donors of São Paulo, Brazil. Vox Sanguinis 74:83–87.

Brown EA, Zhang H, Ping L-H, Lemon SM. 1992. Secondary structure of the 5′ nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. Nucleic Acids Research 20:5041–5045.

Dawson GJ, Schlauder GG, Pilot-Matias TJ, Thiele D, Leary TP, Murphy P, Rosenblatt JE, Simons JN, Martinson FEA, Gutierrez RA, Lentino JR, Pachucki C, Muerhoff AS, Widell A, Tegtmeier G, Desai S, Mushahwar IK. 1996. Prevalence studies of GB Virus-C infection using reverse transcriptase-polymerase chain reaction. J Med Virol 50:97–103.

Fukushi S, Katayama K, Kurihara C, Ishiyama N, Hoshino FB, Ando T, Oya A. 1994. Complete 5′ noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. Biochem Biophys Res Commun 199:425–432.

Heringlake S, Osterkamp S, Trautwein C, Tillman HL, Böker K, Muerhoff S, Mushahwar IK, Hunsmann G, Manns MP. 1996. Association between fulminant hepatic failure and a strain of GBV virus C. Lancet 348:1626–1629.

Hsieh S-Y, Yang P-Y, Chen H-C, Liaw Y-F. 1997. Cloning and characterization of the extreme 5′-terminal sequences of the RNA genomes of GB virus C/hepatitis G virus. Proc Natl Acad Sci USA 94:3206–3210.

Jarvis L M, Davidson F, Hanley JP, Yap PL, Ludlam CA, Simmonds P. 1996. Infection with hepatitis G virus among recipients of plasma products. Lancet. 348:1352–1355.

Katayama K, Kageyama T, Fukushi S, Hoshino FB, Kurihara C, Ishiyama N, Okamura H, Oya A. 1998. Full-length GBV-C/HGV genomes from nine Japanese isolates: characterization by comparative analyses. Arch Virol 143:1063–1075.

Laskus T, Radkowski M, Wang LF, Vargas H, Rakela J. 1997a. Lack of evidence for hepatitis G virus replication in the livers of patients coinfected with hepatitis C and G viruses. J Virol 71:7804–7806.

Laskus T, Wang LF, Radkowski M, Jang SJ, Vargas H, Dodson F, Fung J, Rakela JO. 1997b. Hepatitis G virus infection in American patients with cryptogenic cirrhosis: no evidence for liver replication. J Infect Dis 176:1491–1495.

Leary TP, Muerhoff AS, Simons JN, Pilot-Matias TJ, Erker JC, Chalmers ML, Schlauder GG, Dawson GJ, Desai SM, Mushahwar IK. 1996. Sequence and genomic organization of GBV-C: a novel member of the *Flaviviridae* associated with human non-A-E hepatitis. J Med Virol 48:60–67.

Linnen J, Wages Jr. J, Zhang-Keck Z-Y, Fry KE, Krawczynski KZ, Alter H, Koonin E, Gallagher M, Alter M, Hadziyannis S, Karayiannis P, Fung K, Nakatsuji Y, Shih JW, Young L, Piatak Jr. M, Hoover C, Fernandez J, Chen S, Zou J-C, Morris T, Hyams KC, Ismay S, Lifson JD, Hess G, Foung SKH, Thomas H, Bradley D, Margolis H, Kim JP. 1996. Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. Science 271:505–508.

Muerhoff AS, Leary TP, Simons JN, Pilot-Matias TJ, Dawson GJ, Erker JC, Chalmers ML, Schlauder GG, Desai SM, Mushahwar IK. 1995. Genomic organization of GB viruses A and B: two new members of the *Flaviviridae* associated with GB agent hepatitis. J Virol 69:5621–5630.

Muerhoff AS, Simons JN, Erker JC, Desai SM, Mushahwar IK. 1996a. Identification of conserved nucleotide sequences within the GB virus C 5′-untranslated region: design of PCR primers for detection of viral RNA. J Virol Methods 62:55–62.

Muerhoff AS, Simons JN, Leary TP, Erker JC, Chalmers ML, Pilot-Matias TJ, Dawson GJ, Desai SM, Mushahwar IK. 1996b. Sequence heterogeneity within the 5′-terminal region of the hepatitis GB virus C genome and evidence for genotypes. J Hepatol 25: 379–384.

Muerhoff AS, Smith DB, Leary TP, Erker JC, Desai SM, Mushahwar IK. 1997. Identification of GB virus C variants by phylogenetic analysis of 5′-untranslated and coding region sequences. J Virol 71:6501–6508.

Mukaide M, Mizokami M, Orito E, Ohba K, Nakano T, Ueda R, Hikiji K, Iino S, Shapiro S, Lahat N, Park Y-M, Kim B-S, Oyunsuren T, Rezeig M, Al-Ahdal MN, Lau JYN. 1997. Three different virus C/hepatitis G virus genotypes. FEBS Lett 407:51–58.

Okamoto H, Nakao H, Inoue T, Fukuda M, Kishimoto J, Iizuka H, Tsuda F, Miyakawa Y, Mayumi M. 1997. The entire nucleotide sequences of two GB virus C/hepatitis G virus isolates of distinct genotypes from Japan. J Gen Virol 78:737–745.

Pickering JM, Thomas HC, Karayiannis P. 1997. Predicted secondary structure of the hepatitis G virus and GB virus-A 5′ untranslated regions consistent with an internal ribosome entry site. J Viral Hepat 4:175–184.

Schlueter V, Schmolke S, Stark K, Hess G, Ofenloch-Haehnle B, Engel AM. 1996. Reverse transcription-PCR detection of hepatitis G Virus. J Clin Microbiol 34:2660–2664.

Simons JN, Desai SM, Schultz DE, Lemon SM, Mushahwar IK. 1996. Translation initiation in GB viruses A and C: evidence for internal ribosome entry and implications for genome organization. J Virol 70:6126–6135.

Simons JN, Leary TP, Dawson GJ, Pilot-Matias TJ, Muerhoff AS, Schlauder GG, Desai SM, Mushahwar IK. 1995. Isolation of novel virus-like sequences associated with human hepatitis. Nat Med 6:564–569.

Smuts HE, Kannemeyer J. 1995. Genotyping of hepatitis C virus in South Africa. J Clin Microbiol 33:1679–1681.

Tanaka Y, Mizokami M, Orito E, Ohba K, Kato T, Kondo Y, Mboudjeka I, Zekeng L, Kaptue L, Bikandou B, Pele PM, Takehisa J, Hayami M, Suzuki Y, Gojobori T. 1998. African origin of GB virus C/hepatitis G virus. FEBS Lett 423:143–148.

Tsukiyama-Kohara K, Iizuka N, Kohara M, Nomoto A. 1992. Internal ribosome entry site within hepatitis C virus RNA. J Virol 66:1476–1483.

Tucker TJ, Louw SJ, Robson SC, Isaacs S, Kirsch RE. 1997. High prevalence of GBV-C/hepatitis G virus infection in a rural South African population. J Med Virol 53:225–228.

Yoshiba M, Okamoto H, Mishiro S. 1995. Detection of the GBV-C hepatitis virus genome in serum from patients with fulminant hepatitis of unknown aetiology. Lancet 346:1131–1132.

Zuker M, Jaeger JA, Turner DH. 1991. A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. Nucleic Acids Res 19:2707–2714.